

EP0801344

Biblio Desc

Claims | Page 1

An apparatus for reallocating logical to physical disk devices using

Drawing





a storage controller and method of the same

Patent Number:

-EP0801344, A3

Publication date: 1997-10-15

Inventor(s): YAMAMOTO YASUTOMO (JP); SATOH TAKAO (JP); YAMAMOTO

AKIRA (JP)

Applicant(s):: HITACHI LTD (JP)

Requested Patent:

JP9274544

Application

Number: EP19970105448 19970402 Priority Number(s): JP19960085370 19960408

IPC Classification: G06F3/06

EC Classification: G06F3/06M, G06F11/10M, G06F11/20L

Equivalents: US5956750

Abstract

A storage controller (104) includes that it calculates an access frequency (500) of each logical disk (200); that it selects first logical disk device of which the access frequency exceeds a first predetermined value, the first logical disk device being allocated to a first physical disk device; that it selects a second logical disk device which has the access frequency equal to or less than a second predetermined value, the second logical disk device being allocated to a second physical disk device; and that it reallocates the first and second logical devices to the second

and the first physical disk devices, respectively.



Data supplied from the esp@cenet database - I2

報(A) ধ 盐 华 噩 (2) (2)

(19) 日本国格田(JP)

特開平9-274544

(11)特許出職公開恭母

(43)公開日 平成9年(1997)10月21日

3/09 540 301X 301X 302E 302E 302 302E 302 302E 304 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 320 \$202 \$200 \$202 \$200 \$202 \$200 \$202 \$200 \$202 \$200 \$202 \$200 \$202 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200 \$200	(51) Int C.		## 31 KC++	广内数理每中	FI		技術教示個所
20 2 30 2 E 30 4 N 32 0 7523 - 5 B 12/08 32 0 4 3 4 9 E 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	5	3/06	540			ĸ	0 1
12/08 30.2 30.4			301			30	XI
12/08 32.0 7623-5B 12/08 32.0 32.0 4			302			30	28
12/08 32 0 7623 - 5 B 12/08 32 0 12/08 32 0 12/08 32 0 12/08 32 0 12/08 32 0 12/08 12/			304				NA
特		12/08	N	7623-5B	-		0
中 移職平8-65370 (71) 出職人 00005108 株式会社日立製作所	į				権が開発		19 OL (全14月)
株式会社日立製作所 東京都千代田区神田豊和台四丁目6 銀丸 (72)発明者 山本 康友 神奈川県川崎市康生区王神寺1089母地 式会社日立製作所システム開発研究所 (72)発明者 山本 彰 神奈川県川崎市康生区王神寺1089母地 対金社日立製作所システム開発研究所を (72)発明者 佐藤 孝夫 神奈川県川崎市康生区王神寺1089母地 は72)発明者 佐藤 孝夫 神奈川県川崎市原生区王神寺1089母地 大谷・川県州崎市原生区王神寺1089母地 (72)発明者 佐藤 孝夫	(21) 出国番号		特属平8-85370		(71) 出版人		
山本 度次 特条川県川崎市原生区主導寺1098章地 式会社日立製作所システム開発研究所 山本 第 特条川県川崎市原生区主導寺1099章地 式会社日立製作所システム開発研究所 な会社日立製作所システム開発研究所 佐藤 孝夫 神会川県小田原市国府株280章地 株封 社日立製作所ストレーツシステム事業組 弁理士 有近 幹志郎	(22) (出国日		平成8年(1998)4)	A8B		株式会社日立政作所東京都千代田区神田	表有合四丁目6条地
49.4011.40.11時間 は発生区主事等1033章地 立会社日立製作所システム開発研究所 山本 郡 神奈川県山崎市原生区三澤寺1039章地 文会社日立製作所システム開発研究が 佐藤 孝夫 神参川県小田原市国府株2309章地 株型 社日立製作所ストレーシシステム事業組 弁理士 有近 棒志郎					(72) 完明者	山木東次	
						存金には、自然の関係を対して、対象を行う	X主排型1099排制 株 ステム開発用発用な
					(72) 発明者	14 数	
						特莱川県川崎市康生 式会社日立製作所シ	区王禅寺1039番地 4
					(72)発明者	佐藤 孝夫	
						神会川県小田原市国	府林2880番地 株式会
					1 800 00	社日立教作所ストレ	ーシシステム事業的内
					(74) 代理人	弁理士 有证 养志	題

内部全部状态 (54) [発明の名称]

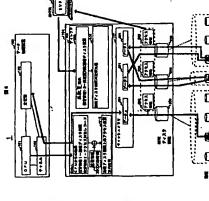
セス情恨500を探取し、そのアクセス情観500をS VPII1を通じて保守員に提示する。保守員の再配置 【解決手段】 各論理ディスク装置200に対するアク 【歌図】 アクセス性能を向上する。 (57) [契約]

指示620があると、指示された2つの論理ディスク技

囧200の間で物理ディスク装置105を配置替えし、

会データを選続的に格扱し直す。

[効果] アクセス頻度の高い論理ディスク装置をより **高速な物理ディスク装置へ再配置することが出来る。シ** よりシーケンシャルアクセス性値の高い物理ディスク装 ーケンシャルアクセスの比耶の高い福理ディスク装置を 置へ再配置することが出来る。



[特許請求の範囲]

【請求項1】 データ処理装置が直接アクセスを行う論 理的記憶装置を実際にデータを記憶する物理的記憶装置 に配置し、前記データ処理装置と前記物理的記憶装置の 間のデータ転送を制御する記憶制御装置において、

予め定めた指標に基づいて前記論理的記憶装置を前記物 理的記憶装置に再配置すると共に再配置先の物理的記憶 装置にデータを連続的に格納する論理的記憶装置再配置 平段を有することを特徴とする記憶制御装置。

【請求項2】 データ処理装置が直接アクセスを行う論 とを対応付け、前記データ処理装置と前記物理的記憶装 前記データ伝送の制御の運用中にデータ処理装置の論理 的記憶装置へのアクセス情報を指標として採取するアク セス情報採取手段と、前記指標に基づいて前記論理的記 記憶装置再配置手段とを有することを特徴とする記憶制 理的記憶装置と実際にデータを記憶する物理的記憶装置 態装置を前記物理的記憶装置に再配置すると共に再配置 先の物理的記憶装置にデータを連続的に格納する論理的 置の間のデータ転送を制御する記憶制御装置において、

情報処理システムに有用である。

[0002]

て、前記アクセス情報が、前記データ処理装置から前記 論理的記憶装置へのアクセス頻度情報を含むことを特徴 「請求項3】 請求項2に記載の記憶制御装置におい とする記憶制御装置。

「請求項4】 請求項2または請求項3に記載の記憶制 **即装置において、前記アクセス情報が、前記データ処理** 装置から前配論理的記憶装置べのアクセスパターン情報 を含むことを特徴とする配憶制御装置。

て、前記指標が、前記論理的記憶装置に求められる信頼 (請求項5) 請求項1に記載の記憶制御装置におい 性であることを特徴とする記憶制御装置。

再配價指示受付手段とを具備したことを特徴とする記憶 【請求項6】 請求項1から請求項5のいずれかに記載 の記憶制御装置において、前配指標を保守員に提示する 指標礎示手段と、保守員からの再配置指示を受け付ける 制御装置。

【請求項7】 請求項1から請求項5のいずれかに記載 指示を受け付ける再配置指示受付手段を具備したことを の記憶制御装置において、データ処理装置からの再配置 特徴とする記憶制御装置。

の記憶制御装置においた、前記指標に基ムいた再配置の 【請求項8】 請求項1から請求項5のいずれかに記載 要否を決定する再配置要否決定手段を具備したことを特 散とする記憶制御装置。 【請求項9】 請求項1から請求項8のいずれかに記載 データ処理装置からのアクセスがあったとき、再配置中 ば再配置先の論理的記憶装置にアクセスさせ、前記アク の記憶制御装置において、再配置中の論理的記憶装置に の論理的記憶装置の再配置完了倒域と再配置未完領域と を徴別し、前記アクセス位置が前配再配置完了領域なら

セス位置が前記再配置未完領域ならば当扱論理的記憶装 屋にアクセスさせるアクセス位歴切替手段をさらに具備 したことを特徴とする記憶制御装配。 [発明の詳細な説明]

[000]

8

し、さらに詳しくは、シーケンシャルアクセスの場合や [発明の属する技術分野] 本発明は、記憶的御装置に関 ランダムアクセルでヒット串が低い場合でもアクセス性 値を向上することが出来る記憶制御装置およびデータの より構成される記憶装置サブシステム、およびその記憶 装置サブシステムとデータ処理装置とにより構成される 特に、本発明は、ディスクアレイ向きの高機能ディスク 装置、その高機能ディスク装置とディスク制御装置とに **佰頼性を向上することが出来る配飽制御装置に関する。** 2

gibson, and R. H. Kartz: A Case for Redundant Arrays o f Inexpensive Disks (RAID), ACM SIGNOD Conference, C SIGMOD」会議において発表された論文「D. Patterson, G. hicago, IL, (June 1988), pp. 109-116」は、ディスクアレ 【従来の技術】シカゴのイリノイ大学で開かれた「ACN イ上のデータ配置に関する技術を開示している。 8

成せずに一旦テンポラリ領域に二直書きし、非同期にバ ディスク装置の一部をディスクキャッシュの如く用いる 技術が開示されている。具体的には、ディスク装置を一 時的にデータを格納するテンポラリ領域と及終的にデー タを沓き込む倒域とに分け、更新データはパリティを生 [0003] また、特開平1-84732号公報では、 リティ生成し、最終領域に書き込む。 52

【0004】一方、電気情報通信学会技術研究報告「D E95-68 (茂木色: Not Wirroring を用いたディス クアレイのディスク故障時の性能評価、1995年12 いる。具体的には、ディスク装置をRAID1構成の簡 分とRAID5構成の部分に分け、ライトアクセスのあ ったデータを優先的にRAID1構成の部分に格納する アクセス頻度の高いデータはRAID1格成の部分に格 納し、アクセス頻度の低いものはRAIDS構成の部分 に格納するように出来る。この技術によれば、記憶容品 の異なる物理ディスク装置やRAIDレベルの異なる物 理ディスク装配を記憶装配サブシステム内で混在させる ことが可能であり、処理ディスク装置内のデータを、そ のアクセス類似やアクセスパターンなどの指指に基づい するRAIDレベルを動的に変更する技術が開示されて 4) 」には、アクセス頻度の近いにより、データを保持 ようにデータの格納位置を動的に変更することにより、 月、電気情報通信学会技程 Vol. 95-No. 407, pp. 19-2 8 33 45

また、アクセス頻度の高いデータを、より高逆な物理デ イスク装置に格納するように、動的に格納位置を変更す て、任意の物理ディスク装置に格納することが出来る。 ることも出来る。なお、RAID1のディスクアレイ

8

-

内閣平9ー274544

RAID1に比べ高い。 [0005]

| つのパリティを作成するため、記憶装置の使用効率は

【発明が解決しようとする認知】上記従来技術では、アクセスするデータ単位でデータの格制位面の変更を行うため、データ処理契配が直接アクセスを行う論理ディスク装配上では近続なデータが、実際にデータを記憶する物理ディスク装配上では非確接となってしまう。このため、一道のデータをリード/ライトするシーケンシャルアクセスの場合、実際には複数データをまとめてリード/ライトできなくなり、アクセス性他の低下を招く同国点がある。

[0006] 一方、上記障告「DE95-68」の従来技術では、ライトの低に、アクセス類度が低いと判断したデータをRAIDI構成の部分からRAIDS構成の部分に移し、空いたRAIDI構成の部分にライトデータを存き込むため、アクセスパターンがランダムアクセルでヒット年が低い場合には、RAIDI構成の部分に戻れてモット年が低い場合には、RAIDI構成の部分に移したデータの多くは再びRAIDS構成の部分に戻れることになる。このため、ヒット年が低い場合、アクセス性能の向上は期待できず、逆にデータを移す処理のオーバヘッドがアクセス性能の低下を引き起こす問題点がある。

【ののの7】また、上記の従来技術では、データの信頼 性の向上については全く考慮されていない問題点があ [0008] そこで、本発明の第1の目的は、シーケンン・トルアクセスの場合やランダムアクセルでヒット申が低い場合でも、アクセス性能を向上することが出来る記憶制御装配を提供することにある。また、本発明の第2の目的は、データの信頼性を向上することが出来る記憶別砂袋配を提供することにある。

[0000]

(現伍を解決するための手段) 第1の観点では、本発明 45 は、データ処理装配が直接アクセスを行う論理的記憶装 許を実際にデータを記憶する物理的記憶装置に配置し、 相記データ処理装置と加記物理的記憶装置の面のデータ 転送を納抑する記憶的御装置において、予め定めた指標 に表づいて前記論理的記憶装置を削記物理的記憶装置。50

再配置すると共に再配置先の物理的記憶装型にデータを 連接的に格納する論理的記憶装置再配置手段を有するこ とを特徴とする記憶制御装置を提供する。上記第1の観 点による記憶制御装置を提供する。上記第1の観 点による記憶制御装置では、アクセスするデータ単位で 5 データの格制位置の変更を行うのではなく、論理的記憶 装置を単位として物理的記憶装置にデータを連接的に格納 する、従って、シーケンシャルアクトエスの場合でも、ア クセス性能を向上することが出来る。また、ライトの度 のにデータの格特位置の変更を行うのではなく、チン定め た指域に基へいて加足再足型を行うから、ランダムアク せんでとット車が低い場合でも、アクエス化能を向上す ることが出来る。

い、且つ、再配置先の物理的記憶装置にデータを連続的 に格納する。従って、ジーケンジャルアクセスの場合で も、アクセス性能を向上することが出来る。また、ライ トの度にデーダの格制位置の変更を行うのではなく、ア クセス情報を採取し、それを統計的に利用して前配再配 置を行うから、ランダムアクセルでヒット率が低い場合 でも、アクセス性能を向上することが出来る。

35 【0011】第3の異点では、本発明は、上記構成の記憶制御装置において、前記アクセス情報が、前記データ 処理装置から前記論理的記憶装置へのアクセス頻模が、自己データ 処理装置から立てを特徴とする記憶制御装置へのアクセス頻度の高 第3の製点による記憶制御装置では、アクセス頻度の高 40 い論理的記憶装置をより高速な物理的記憶装置へ再配置 することが出来る。従って、アクセス性能を向上することが出来る。従って、アクセス性能を向上することが出来る。

【0012】第4の報点では、本発明は、上記構成の記憶制が装置において、前記アクセス情報が、前記データ処理装置から前記録単的記憶装置へのアクセスパターン情報を含むことを特徴とする記憶制御装置では、シーケンシャルアクセスの比単の高い論理的記憶装置をよりシーケンシャルアクセス性能の高い論理的記憶装置をよりシーケンシャルアクセス性能の高い物理的記憶装置をよりシーケンジャルアクセス性能の高い物理的記憶装置をよりシーケンジャルアクセス性能の高い物理的記憶装置を上りシーテンジャルアクセス性能の高い物理的記憶装置を上りたことが出来る。従って、アクセス性能を向上することが

出来る。

[0013] 第5の製点では、本発明は、上記構成の記憶制御装置において、前記指填が、前記舞曲的記憶装置に次められる信頼性であることを特徴とする記憶制御装置を提供する。上記第5の製点による記憶制御装置では、信頼性が高いことが求められる論理的記憶装置でよい、信頼性の高い物理的記憶装置へ与配置することが出来る。後つて、データの信頼性を向上することが出来る。後つて、データの信頼体を向上することが出来る。後の現点では、本発明は、上記構成の記憶制御装置において、前記指域を保存員に総示する指域を示す及と、保守員かの再配置相示を受け付ける再配置相示を決して、ことを特徴とする記憶制御装置では、保守員が再配置指示を入力できるため、非常に柔軟に前配再配置を行うことが出来る。

[0015]第7の観点では、本発明は、上記構成の記憶例御装屋において、データ処理数位からの再配配指示を受け付ける再配配指示受付手段を具備したことを特徴とする記憶制御装置を提供する。上記第7の観点による記憶制御装置では、データ処理装置が再配置指示を入力できるため、保守員では判断不可能な高度の条件下で前記再配置を行うことが出来る。

[0016] 第8の観点では、本発明は、上記構成の記憶制御装置において、前記指環に基づいて再配置の要否を決定する再配度の要否決定手段を具備したことを特徴とする記憶制御装置を提供する。上記第8の観点による記憶制御装置では、記憶制御装置が再配置指示を自己決定するため、保守員やデータ処理装置に負担をかけなくてある。

【0017】第9の観点では、本発明は、上記構成の記憶制御装団において、再配置中の論理的記憶装置にデータ処理装置からのアクセスがあったとき、再配置中の論理的記憶装置にデータ処理装置からのアクセスがあったとき、再配置中の論理的記憶装置にデースをせるアクセス位置が静記程でに対するとは、前記アクセス位置が静記程でに、前記度が開発を提供する。上記第9の観点にアクセスさせるアクセス位置切替手段を具備したことを特徴とする記憶制御装置では、再配置中の路理的記憶装置の再配置完了領域と有配置未完領域とを確別し、データ処理装置からのアクセス位置を切り替えるから、データ処理装置と物理的記憶装置の両のデータ転送を運用中に再配置を行うことが出来る。

[0018]

【発明の実施の形態】以下、本発明の実施形態を説明する。なお、これにより本発明が限定されるものではな

[0019] - 第1の実施形態-

第1の実施形態は、各種理ディスク装置のアクセス情報 を記憶制御装置で探覧し、SVP(サービスプロセッ

サ)を適じて保午員に提示し、このアクセス情報に基づく保午員の再配配指示により、韓理ディスク装型の物理ディスク装型の物理ディスク装型への再出配を行うものである。

43

[0020] 図1は、本発明の第1の実施形態にかかる 5 記憶制御装置を含む荷輪処理システムのブロック図であ 5 この情報処理システム1は、データ処理装図100 と、記憶制御装図104と、1台以上の物理ディスク装 歴105と、SVP111とを接続してなっている。 [0021] 前記データ処理装図100は、CPU10

(0022] 前記記憶却御装置104は、1つ以上のディレクタ106と、キャッシュメモリ107と、ディレクトリ108と、不得発性メモリ109と、不得発性メニリ 108と、不得発性メニリロのと、不得発性メニリロのと、不得発性メニリアが3。前記ディレクタ106は、データ処理装置100のディネル103と前記キャッシュメモリ107と物理ディスク装置105の間のデータ配送、データ処理装置100のディネル103と前記キャッシュメモリ107と物理ディスク装置105の間のデータ配送を行う。前記キャッシュメモリ107には、物理ディスク装置105の中のアクセスが低の高い、レクを104に対したのよいを104を104によった。101に前に、1024

25 レクタ106が実行する。ロードするデータの具体的は、データ処理装置100のCPU101のプクセス対象データや、このアクセス対象データと物理ディスク数 図105上の格特位図が近いデータ等である。前記ディレクトリ108は、前記キャッシュメモリ107の管理 30 情報を格辞する。前記不得発性メモリ109は、前記キャッシュメモリ107と回域に、物理ディスク装成105の中のアクセス頻度の高いデータをロードしておく。前記不得発性メモリ管理情報を指針する。前記海連物里対応信号 報300は、各籍理ディスク装成 (図2の200) が配

前配不得発性メモリ管理情報110は、前記不得発性メモリ109の管理情報も右続する。前記過程物理対応信 モリ109の管理情報を格許する。前記過程物理対応信 電300は、各過程ディスク装配105上の位置およびか物理ディスク装配105上の位置およびか物理ディスク装配105上の位置を用いて、データ処理装配100のCPU101のアクセス 40 対象データの物理ディスク装配105上の格特図域の算出などを行う。前記論理ディスク装配105上の格特図域の算出などを行う。前記論理ディスク装配105上の格特図域の算出などを行う。前記論理ディスク装配105上の格特図域の算出などを行う。前記論理ディスク装配(図20200)のアクセス可容等の状態を示す。前記アクセス格限500は、各過程ディスク装置(図20200)のアクセスが度くスクセスバタ

45 一ンなどの情報である。 (0023) 福理物理対応情報300と福田ディスク信 報400は、電級断などによる消失を訪ぐために不得我 の媒体に記録する。

[0024] 前記物理ディスク装置105は、データを 50 記録する媒体と、記録されたデータを認み書きする装置

ç

からは成される。

[0025] 前記3~P111は、プラセス情報500の保守員への投示や、保守員からの再配置指示620の人力の受け付けを行う。また、保守員からの情報処理・ステム1~の指示の発信や、情報処理・ステム1の確否 の状態等の保守員への起示を行う。

1が直接アクセスする見掛け上のディスク装置で、アク セス対象データが実際に格納される物理ディスク装置 1 タは、シーケンシャルアクセスを考慮して、物理ディス ク装置105上に連接的に配置されている。論理ディス ク装置200のデータが配置されている物理ディスク装 R105がディスクアレイ構成の場合、核論理ディスク な肌200と対応する。この論理ディスク装置200と 物理ディスク装取 105の対応は前記論理物理対応情報 05の領域内のデータ格納位限202を求め、データ転 【0026】因2は、諸理ディスク装置200と物理デ イスク装成105の因迹を表わした因である。 論理ディ スク装置200は、データ処理装置100のCPU10 05と対応している。 确理ディスク装置200上のデー 5。また、物理ディスク装置105の容量が論理ディス ク装限200より大きく、複数の論理ディスク装置のデ ータを1台の物理ディスク装成105に格納できる場合 300で管理される。例えば、データ処理装置100の CPU101が簪埕ディスク装図200のデータ201 位300に基づき舗理ディスク装置200に対応する物 には、該物理ディスク装置105は複数の論理ディスク をリードする時、記憶制御装置104で論理物理対応情 **咒ディスク装置105を求め、その物理ディスク装置1** 技匠200は複数の物理ディスク装置105と対応す

[0027] 図3は、緯理物理対応情報300を表わした図である。 編理物理対応情報300は、韓理ディスグ構成情報320とから構成される。 前記論理ディスク構成情報320とから構成される。 前記論理ディスク構成情報310は、各籍理ディスク集成200から対応する物理ディスク装配105上の領域に関する情報であり、論理ディスク装置20に、各物理ディスク構成情報320は、各物理ディスク構成情報320は、各物理ディスク構成に関立されている論理ディスク装成200に関する情報で、物理ディスク装成105に配置されている論理ディスク装成200に関する情報で、物理ディスク装成105に配置されている論理ディスク装成200に関する情報で、物理ディスク装成105に配置されている論理ディスク装成200に関する情報で、物理ディスク装成105に

[0028] 前記編程ディスク構成情報310は、物理ディスク装成グループ311, RAID構成312および開始(位312の記念・福理ディスク装成2000数 45だけ たしている。前記物理ディスク装成グループ311は、当路温電ディスク装成200が配置されている物理ディスク装成105を示す情報である。前記RAID構成312は、前記物理ディスク装成グループ311のRAIDにいたルを示す。前記開始位展313は、当該韓國・50

ディスク装置200が物理ディスク装置105上で配置 されている先頭位置を示す。 [0029] 前記物理ディスク構成情報320は、結理ディスク装置10 ディスク装置グループ321を、物理ディスク装置10 5の数だけ有している。前記論理ディスク装置グループ321は、当該物理ディスク装置105に配置されている結理ディスク装置200を示す。

[0030] 図4は、簡型ディスク情報400を抜むした図である。 論型ディスク情報400は、論理ディスク状態401と再配置ディンタ402とを、論理ディスクスク装図200数数に付在している。 前記論理ディスク状態401は、「正常」 閉塞」「コオーマット中」

「再配度中」などの論理ディスク装配200の状態を装わす。前記再配配系アポインタ402は、前記論理ディスク状態401が「再配置中」の時のみ有効な情報で、当球論理ディスク装置200の再配置処理を完了している領域の大の位置すなわら当該議理ディスク装置200・米に再配置処理を終えていない関域の先頭位置を示していまって、

す。「再配置中」におけるデータアクセス時、再配置売20 Tポインタ402 草りも前の領域へのアクセスの場合には、再配度後の物理ディスク装配105ヘアクセスしたければならない。一方、再配置完了ポインタ402以後の領域へのアクセスの場合には、再配置前の物理ディスク装配105ヘアクセスしなければならない。

55 [0031] 図5は、アクセス情報500を変わしてい 5。アクセス情報500は、アクセス頻度情報501と アクセスパケーン情報502とを、諸理ディスク装置2 00の数だけ有している。このアクセス情報500は、 配協制御装置104、データ処理装置100。SVP1 30 11のいずれからも参照することが出来る。前記アクセス 外に情報501は、単位時間あたりの当該過程ディスク が概定情報501は、母位時間あたりの当該過程ディスク が概定情報501は、各強理ディスク装置200の中で アクナスが振り落しませます。このアクセスが成に関係501は、各強理ディスク装置2000中で

11のいずれからも部別することが出来る。前記アクセス類反情報501は、単位時間あたりの当該協理ディスク装置200つのアクセス回数を管理する。このアクセス類位権601は、各種程ディスク装置200の中でアクセス数度201は、各種程ディスク装置200の中でアクセスが優200のシーケンシャルアクセスとデングエアクセスの割合を管理する。このアクセスとラングエアクセスの割合を管理する。このアクセスバクーンが開発502は、シーケンシャルアクセスが多く、よりシーケンシャルを能の高い物理ディスク装置200を求める指摘として用いる。

[0032] 次に、記憶制御装置104の動作を説明する。図6は、記憶制御装置104の動作を詳細に表むした図である。まず、リード/ライト処理時の動作について説明する。ディレクタ106は、通常リード/ライト処理を実行する際、CPU101からチャネル103を程由してCPUからの指示600を受け取る。このCPUからの指示600は、リード(またはライト)対象のレコードが記憶されている論理ディスク装置200を指定する指定情報1と、リード(生たはライト)対象のレ

コードが記憶されている論理ディスク装包200内の位置 (トラック, セクタ, レコード) を指定する指定情報 2とを含んでいる。ディレクタ106は、物理ディスク 装置上のアクセス位置算出処理 (610)で、前記CP Uからの指示60と論理物理対応情報300とを用いて、物理ディスク装置105上でのアクセス位置を算出する。この物理ディスク装置アクドス位置算出処理 (610)については図8を参照して後で詳述する。その 後、たとえばフード処理では、算出した効理ディスク装領105上のデータ格納位置202が一分をキャッショメモリ101にからにボジル上げてデータ201とし、その 数み上げたデータ201をディネル103を通じて主記 (6102に伝送する。

(0033)次に、プウセス情報500の採む処理について説明する。CPU101からのリード/ライト処理のアクセス時に、ディレクタ106は、アクセス対象論理ディスク数数00のアクセス情報500を更新する。アクセス類度情報501の疾動は、例えば、アクセス類度情報501の疾動は、例えば、アクセスが同じがカンクをカウントアップしていき、一定時間または一定回数のアクセスがは、例えば、アクセス内のアンクにシーケンシャルアクセス回転に対していき、一定時間または一定回数のアクセスが、多ーンを判定する。

するデータ低に決定される。

[0034]次に、再配度指示620を投明する。保守員は、SVP111を通じて提示されたアクセス情報500を参照して、各種理プィスク装置200の再配置の必要性を検討する。この検討の結果、再配置を決定した配置的非決定。この再配置指示620は、再配置対象の論理ディスク装置200を20は、再配置対象の論理ディスク装置200を20は、再配置対象の論理ディスク装置200を20は、再配置対象の論理ディスク装置200を20指定する指示情報1-2からなる。保守員が行う検討の内容は、後述する第3の実施形態で図10を無限して説明する論理ディスク装置再配置要否決定処理(910)と同様である。

6.20を契けて、指定された2つの韓理ディスク数度2000回で論理ディスク数度2位000回で論理ディスク数度再配置処理(630)を行う。図7は、韓理ディスク数度再配置処理(630)を行う。図7は、韓理ディスク数度再配置処理(300処理フロー図である。ステップ700では、韓理ディスク数位200%理ディスク数位200%理ディスク数位200%理ディスク数位200%理ディスク数位200%更近2012がインタ402を各論理ディスク数位200の再配位に初期化する。ステップ702では、韓理ディスタ数で200の方面位置に初期化する。ステップ702では、韓理ディスク数

☆200の再配型会Tポインタ402をチェックし、金額板の再配置が完了していなければステップ 703 〜進み、完了していればステップ 707 〜進む。

•3

[0036] ステップ703では、再配配売りポインタ05 402が示すデータ位置から再配配処理の1回の処理単位分のデータに対して物理ディスク装配105からキャッシュメモリ107上へのデータ転送を行う。ここで、1回の処理単位分のデータ点は、再配限対象の2つの設理ディスク装配20の元長データ1つに対応する・ストスク装置200とRA1D1の論理ディスク装置200とRA1D1の論理ディスク装置200の元氏データ1つに対応するデータ は、RA1D5の論理ディスク装置200元度データ1つに対応するデータ はは、RA1D5の論理ディスク装置200元度データ1つに対応するデータ は、RA1D5の論理ディスク装置200元度データ 1つに対応するデータ は、RA1D5の論理ディスク装置200元度データ 1つに対応するデータ 1つに対応

(0031) ステップ 704では、再配配対象の各路四 ディスク装配 200の再配配供売組甲ディスク装配 200 がパリティを有するRAIDレベルのものである場合、キャッシュメモリ107上の再配配対象の1回の処理単位分のデータ 201に対してペリティを生成する。ステップ 705では、キャッシュメモリ107上の前配対 象の1回の処理単位分のデータ 201および前記ステック 704で作成したパリティを、用配置先の物理ディスク装配 105~普音込む。ステップ 706では、1回の処理単位分だけ形配包完了ポインタ 402を進める。そして、前記ステップ 702に戻る。

[0038] なお、上記ステップ703,704において、データおよびパリティは、不仰受性メモリ109にも転送して二重化し、キャッシュの官によるデータ消失を防ぐ。この理由は、上記ステップ705での許き込み時に、例えば、第1の論理ディスク装置2005年20 論理ディスク装置2000データのうち、第1の論理ディスク装置2000データを物理ディスク装置10

38 イスク装図2000データを物理ディスク装図105 (元は第2の路理ディスク装図200に配置されていた 物理ディスク装図105) へ音き込んだ段所で協巧によ りキャッシュメモリ107上のデータがアクセス不能に なったとすると、容き込みが終了してない第2の論理デ 40 イスク装図200のデータが消失するからである (元は 第2の論理ディスク装図200に配置されていた物理ディスク装図105には、上記のように第10論程ディスク が変図200のデータが上むきされてしまっている)。 [0039] ステップ707では、論理物理対応情報3

【003名】次に、論理ディスク装置再配置処理(63

 50 [0040] 次に、物理ディスク装置アクセス位置算出

.

•)

「再配置中」ならばステップ801に進み、「再配置中 処理(610)を説明する。図8は、物理ディスク装置 ステップ800では、論理ディスク情報400のうちの アクセス対象論理ディスク技假200の論理ディスク状 アクセス位配算出処理部610の処理フロー図である。 堕101が「再配置中」であるか否かをチェックし、

配置完了ポインタ402とアクセスデータ位置とを比較 し、アクセスデータ位置が再配置完了ポインタ402の 情寸位別以後ならばステップ802に進み、アクセスデ - 夕位氏が再配因完了ポインタ402の指す位置より前 【0041】ステップ801では、論理ディスク情報4 0 0のうちのアクセス対象論理ディスク装配200の再 で」なければステップ803に進む。 ならばステップ803に進む。

[0042] ステップ802では、当該論理ディスク装 **M200の再配円先の論理ディスク装置200をアクセ** ス対象にする。そして、ステップ804〜遊む。

[0043] ステップ803では、当該論理ディスク装 聞200をアクセス対象とする。

[0044] ステップ804では、アクセス対象の諸型 ディスク技配 200に対応した物理ディスク装置 105 上でのアクセス位配を、論理物理対応情報300を用い て草川する。

ステム1および記憶制御装置104によれば、アクセス 情報500に基づく保守員の判断により、アクセス頻度 の高い論理ディスク装置をより高速な物理ディスク装置 とが出来る。従って、アクセス性能を向上することが出 【0045】以上の第1の実施形態にかかる情報処理シ セスの比単の再い福理ディスク装置をよりシーケンシャ ルアクセス性能の高い物理ディスク装置へ再配置するこ へ再配贷することが出来る。また、シーケンシャルアク

[0046] -第2の実施形態-

上記第1の実施形態を変形して、記憶制御装置1.04か 御装页104に再配配指示(620相当)を出すように し、データ処理装取100が再配置要否を決定し記憶制 らアクセス情報500をデータ処理装置100に提示 してもよい。

33

[0047] - 第3の実施形態-

第3の実施形態は、再配置指示をSVP111やデータ 処理技能100から受けるのではなく、配飽制御装置1 0.4が自己決定するものである。

[0048] 図9は、記憶制御装置104の動作を詳細 は、論理ディスク再配置要否決定処理部910が再配置 に表わした図である。第1の実施形態(図6)との違い 指示620を出すことである。

が一定周期で各番理ディスク装置200のアクセス情報 [0049] 図10は、上記論理ディスク再配置要否決 ク再配限要否決定処理(910)は、ディレクタ106 定処理部910の処理フロー図である。この論理ディス

上記第1一第3の実施形態を変形して、アクセス情報5 00に代えて又は加えて、論理ディスク装置200に要 50 水される信頼性を再配置処理要否決定の指標に用いても

園(以下、これを第1候補論理ディスク装置という)2 ス情報500のアクセス頻度情報501を参照し、アク セス頻度が規定値を組え且つ配置されている物理ディス ク装図 1 0 5 が比較的低速なものである論理ディスク装 装置200があればステップ1001〜進み、なければ 500を検査して行う。 ステップ1000では、アクセ 00があるか否かをチェックし、該当する論理ディスク ステップ1005〜進む。

[0050] ステップ1001では、前記第1候補論理 ディスク装置200のアクセスパターン情報502を畚 **照し、シーケンシャルアクセスの比串が規定値以上であ** るか否かをチェックし、規定値以上でなければステップ 1002~進み、規定値以上であればステップ1004 2

[0051] ステップ1002では、前記第1候補論理 度情報 5 0 1 を参照し、アクセス頻度が規定値以下の論 ディスク装置200より高速な物理ディスク装置105 に配置されている論理ディスク装置200のアクセス頻 理ディスク装置 (以下、これを第2候補論理ディスク装 20 置という) 200 崩あるか否かをチェックし、あればス テップ1003~進み、なければステップ1005~進 2

【0052】ステップ1003では、前記第1候補論理 ディスク装置200と前記第2候補論理ディスク装置2 し、再配置指示620を出す。そして、処理を終了す 00の間で再配置処理 (630) が必要であると決定

[0053] ステップ1004では、前配第1候補臨理 ディスク数配200よりシーケンシャル性能の高い物理 ディスク装置105に配置されている論理ディスク装置 2000アクセスパターン信仰502を参照し、ツーケ ソシャルアクセスの比率が規定値以下の福理ディスク装 ত (以下、これを第2候補論理ディスク装置という) 2 0.0があるか否かをチェックし、あれば前記ステップ1 003~進み、なければ前記ステップ1002~進む。 200の再配限処理 (630) は不要であると決定す る。そして、処理を終了する。

情報500に基づいて自動的に、アクセス頻度の高い論 【0055】以上の第3の実施形態にかかる情報処理シ ステム1および配億制御装置104によれば、アクセス 理ディスク装置をより高速な物理ディスク装置へ再配置 **することが出来る。また、シーケンシャルアクセスの比 串の高い倫理ディスク装置をよりシーケンシャルアクセ** ス性能の高い物理ディスク装置へ再配置することが出来 5. 従って、アクセス性能を向上することが出来る。 [0056] -第4の実施形態-\$ 45

…情報処理システム 100…データ処理装団 104…配物的物数型 106...ディレクタ 103…チャネル 101...CPU 102…主配包 [作号の説明] 8 【図1】本発明の第1の実施形態にかかる記憶制御装置 ソシャルアクセスの場合やランダムアクセルでヒット母 よい。信頼性を指標に用いれば、論理ディスク装置20 5。また、本発明の記憶制御装置によれば、データの信 【発明の効果】本発明の記憶制御装置によれば、シーケ が低い場合でも、アクセス性能を向上することが出来 0 上のデータの信頼性を向上させることが出来る。 質性を向上することが出来る。 [図面の簡単な説明] [0057]

110…不揮発性メモリ管理情報 108…キャッシュディレクトリ 105…物理ディスク装置 10 107…キャッシュメモリ 109…不符配在メモリ 【図2】論理ディスク装置と物理ディスク装置との対応 【図3】 論理物理対応情報の構成例示図である。 を合む情報処理システムのプロック図である。

1111...SVP

15 200…福理ディスク装囚 300…論理物理対応信報 202…データ格制位因 201 -- 7-5 [図6] 本発明の第1の実施形態における記憶制御装置

[図4] 陰根ディスク情報の構成例示図である。

関係の説明図である。

【図5】 アクセス情報の構成例示図である。

の製作やドナノロック図わせる。

400…福理ディスク信報 600…CPUからの指示 500…アクセス信報 【図7】 論理ディスク装置再配置処理部の処理フロー図 [図8] 物理ディスク装置アクセス位置算出処理部の処

610…物理ディスク装置上のアクセス位置算出処理部 620…指示情報

630…論理ディスク装置再配置処理部 [図9] 本発明の第3の実施形態における記憶制御装置

[図10] 路理ディスク装置再配置要否決定処理師の処 25 910…路理ディスク再配置要否決定処理部

の動作を示すプロック図である。

型フロー図である。

埋フロー図である。

[図3]

(K)

福理物理刘応情報

・路理ディスク様成情報 310

60 - 60 **物理ディスク核電グルー** R A i D 様成 情理ティスク 転置の数だけ 用意

物理ディスク構成情報 320

物理ティスク { 【論理ディスク装置グループ 路間の数だけ { 用監

8 1

[図2] 2 E

[图7] 1

SVP

チャネル

チャネル

記住制御城區

アータ処理技団

_ E

情假处理シスナム

[図1]

共同語

OPO

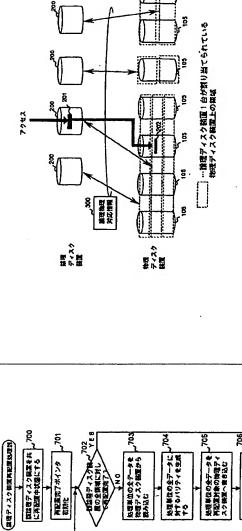


图2

<u>8</u>

アクセス情報

- 01 -

路理ティスク 接限の数だけ | 麻理ディスク 用意

4

論理ディスク債権

[図4]

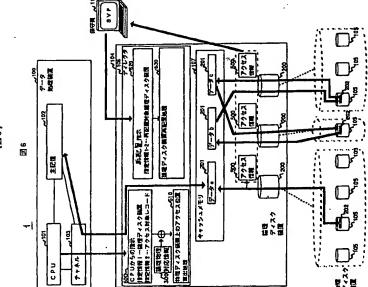
物理ディスク雑画

理ディスク知面

910

(物理ディスク経送アクセス位属算出部

[図8]



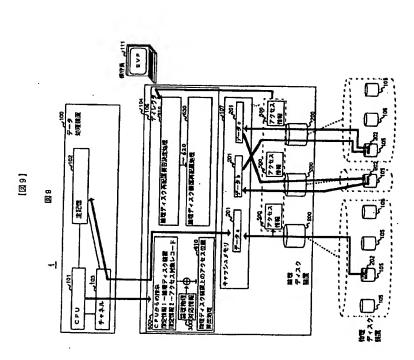
[86]

(登理ディスク価医再配置終否決定処理語)

읡

[國] 0]

0 |



1,1003 石配置不器と特定

- 13 -

:. ..